The EMBO Journal vol.9 no.1 pp.161 – 169, 1990

Characterization of the sequence-specific interaction of mouse c-*myb* protein with DNA

K.M.Howe, C.F.L.Reakes and R.J.Watson

Imperial Cancer Research Fund Laboratories, St Bartholomew's Hospital, Dominion House, Bartholomew Close, London EC1A 7BE, UK

Communicated by I.M.Kerr

We have examined parameters that affect sequence-specific interactions of the mouse c-*myb* protein with DNA oligomers containing the Myb-binding motif (CA/CGTTPu). Complexes formed between these oligomers and *in vitro* translated c-*myb* proteins were analysed by electrophoresis on non-denaturing polyacrylamide gels using the mobility-shift assay. By progressive truncation of c-*myb* coding sequences it was demonstrated that amino acids downstream of a region of three imperfect 51–52 residue repeats (designated R1, R2 and R3), which are located close to the amino terminus of the protein, had no qualitative or quantitative effect on the ability to interact specifically with this DNA motif. However, removal of only five amino acids of the R3 repeat completely abolished this activity. The contribution of individual DNA-binding domain repeats to this interaction was investigated by precisely deleting each individually: it was demonstrated that a combination of R2 and R3 was absolutely required for complex formation while the R1 repeat was completely dispensible. c-*myb* proteins showed quantitatively greater interaction with oligomers containing duplicated rather than single Myb-binding motifs, in particular where these were arranged in tandem. Moreover, it was observed that c-*myb* protein interacted with these tandem motifs as a monomer. These findings imply that a single protein subunit straddles adjacent binding sites and the implications for c-*myb* activity are discussed.

Key words: c-*myb*/DNA-binding protein/Myb-binding motif

Introduction

The c-*myb* gene has been implicated in the induction of a number of haemopoietic malignancies in experimental animals. For example, transduction of c-*myb* sequences by the avian retroviruses AMV and E26 was associated with leukaemias of immature myeloid phenotype (Bishop and Varmus, 1985), while insertion of retroviruses into this locus has been observed in certain murine myeloid leukaemias (Mushinski *et al.*, 1983; Weinstein *et al.*, 1986) and avian B cell lymphomas (Kanter *et al.*, 1988; Pizer and Humphries, 1989). Typically, these induction events involved disruption of c-*myb* coding sequences, resulting in expression of truncated analogues of the 75 000 mol. wt c-*myb* protein (Klempnauer *et al.*, 1983; Boyle *et al.*, 1986). Thus, AMV encodes a 45 000 mol. wt protein comprising six amino acids specified by the retrovirus *gag* gene fused to c-*myb* sequences truncated at both the amino and carboxy termini (Klempnauer *et al.*, 1983). In some contrast, retrovirus insertion appears to result in expression of genes that encode proteins singly truncated at either the amino or carboxy termini (Lavu and Reddy, 1986; Shen-Ong *et al.*, 1986; Kanter *et al.*, 1988; Pizer and Humphries, 1989).

Both authentic and truncated *myb* proteins are located in the cell nucleus, where they are associated with chromatin (Boyle *et al.*, 1984; Klempnauer *et al.*, 1984; Moelling *et al.*, 1985; Klempnauer and Sippel, 1986). Consistent with this subcellular localization, it has been found that these proteins may directly and non-specifically interact with DNA *in vitro* (Moelling *et al.*, 1985; Klempnauer and Sippel, 1986; Bading *et al.*, 1987). The region of the v-*myb* protein that is required for this non-specific interaction has been localized to a partially reiterated sequence at the amino terminus (Klempnauer and Sippel, 1987; Ibanez *et al.*, 1988). The equivalent region of c-*myb* comprises three imperfect repeats of 51–52 amino acids. Interestingly, a major part of the first of these repeats is deleted in the AMV and E26 v-*myb* genes and is similarly truncated by insertional mutagenesis in most murine myeloid leukaemias in which c-*myb* is implicated. This reiterated structure is strongly conserved through evolution; the mouse and human protein sequences are identical in this region (Gonda *et al.*, 1985; Bender and Kuehl, 1986; Majello *et al.*, 1986; Slamon *et al.*, 1986) and chicken c-*myb* differs by only three amino acid substitutions (Rosson and Reddy, 1986), all of which are located in the first repeat. Furthermore, this arrangement is conserved in a *Drosophila* homologue (Katzen *et al.*, 1985; Peters *et al.*, 1987) and in two human c-*myb*-related genes (Nomura *et al.*, 1988), even though there is little conservation of amino acid sequence outside of this region. Only in the cI gene, a homologue of *myb* in maize, is the three repeat structure not retained, the first repeat being absent in this instance (Paz-Ares *et al.*, 1987).

The extent of conservation of the reiterated structure implies a critical role in determining DNA binding of this protein with specific targets. Indeed, evidence for sequence specificity of this interaction was obtained using purified human c-*myb* protein in a filter retention assay (Bading *et al.*, 1987). Retention of a single 2.0 kb DNA restriction fragment from bacteriophage λ was observed, and it was later demonstrated that an 81 bp subfragment could be recognized by purified v-*myb* protein in a mobility-shift assay (Bading, 1988). From the gel mobilities of the complexes formed it was suggested that v-*myb* protein bound DNA both as a monomer and as a dimer, but that strong interactions were obtained only with a dimer. Recently, a consensus sequence (PyAACG/TG) involved in specific interaction of v-*myb* protein with DNA has been identified using an assay in which cloned random fragments of chicken DNA were screened with bacterially expressed v-*myb* protein bound to filters (Biedenkapp *et al.*, 1988). Typically, plasmids selected in this way had chicken DNA inserts containing multiple

copies of the hexameric Myb-binding motif, many of which were arranged in close apposition and which were shown to bind the v-*myb* protein in DNase-footprinting experiments. It is notable that the 81 bp λ DNA fragment that binds v-*myb* protein contains two perfect copies of the consensus Myb-binding motif spaced two nucleotides apart. We report here on experiments in which oligomers based on this λ DNA sequence were used to examine parameters affecting DNA−protein complex formation. In particular, we wished to examine the role played in this interaction by the three c-*myb* protein repeats previously implicated in generalized DNA binding.

Results

DNA binding by in vitro-translated c-myb protein

Full-length mouse c-*myb* protein was made by *in vitro* translation of a synthetic mRNA, transcribed from a cDNA-containing plasmid, pT7βmyb, with T7 RNA polymerase. To determine whether the c-*myb* protein binds DNA with the same sequence specificity as reported for the truncated v-*myb* product (Biedenkapp *et al.*, 1988), mobility-shift assays were employed. The *in vitro*-translated c-*myb* protein without prior purification was incubated with a synthetic double-stranded DNA oligomer (OHL34/35) that contains two consensus hexameric Myb-binding motifs (CAGTTA) spaced four nucleotides apart (see legend to Figure 1). The sequence of this oligomer was based upon that of an 81 bp *Sph*I−*Hin*fI fragment of bacteriophage λ DNA shown previously to bind v-*myb* protein (Bading, 1988). Upon incubation with reticulocyte lysates containing c-*myb* protein, a complex was formed which resulted in a shift in the mobility of the oligomer on non-denaturing polyacrylamide gel electrophoresis (Figure 1). The formation of this complex was resistant to competition by a large excess (2 μg) of poly(dI-dC). However, this interaction could be competed by co-incubation with modest amounts (10 ng) of unlabelled OHL34/35 oligomer, but not by similar or greater amounts of an unrelated oligomer (OHL7/12) that lacks any Myb-binding motifs. These results therefore point to a specific interaction between c-*myb* protein and a DNA fragment containing the consensus binding sites.

c-myb protein sequences involved in DNA binding

We next examined the effects of carboxy-terminal truncation of c-*myb* protein on DNA binding. The motives for this were threefold. First, if the mobility of the protein−DNA complex was found to be dependent on the size of the *in vitro*-translated product, this would provide further evidence that the mobility shift observed above reflected an interaction between the oligomer and Myb. Second, we wished to establish the limits of the protein domain implicated in specific DNA binding and, in particular, to determine whether this corresponded to the amino-terminal v-*myb* region shown previously to have a general affinity for DNA (Klempnauer and Sippel, 1987). Third, it was of interest to determine whether other features of the c-*myb* protein sequence outside this region had any influence, either at a qualitative or quantitative level, on this activity. In this regard, it is notable that c-*myb* contains a short sequence comprising leucines and other hydrophobic residues arrayed at regular seven amino acid intervals (Met/Ile/Leu/Leu/Leu, amino acids 375−403) which may be analogous to a

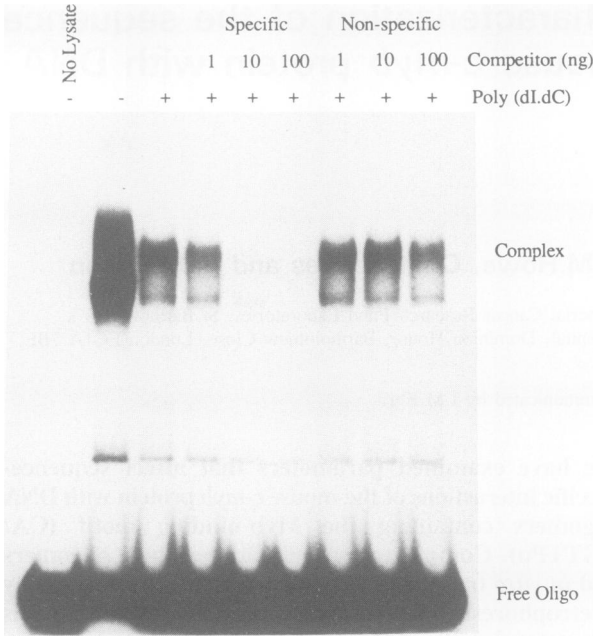

Fig. 1. Mobility-shift analyses of c-*myb* protein−DNA complexes. 250 pg of the ^{32}P-labelled double-stranded DNA oligomer OHL34/35 (5′-CTAGAATTGACAGTTAATAGCAGTTAATTTT, see also legend to Figure 6) was incubated in the absence or presence of reticulocyte lysates containing *in vitro*-translated c-*myb* protein. Some reactions included 2 μg poly(dI-dC) and varying amounts (1−100 ng) of either a specific DNA competitor (OHL34/35) or a non-specific competitor (OHL7/12; 5′-TATGAATGTTACATTCTATCAAGGAAACAAACC-GTGCAAAGAAACGAAAAAGAACACTGG) that lacks the Myb-binding motif. DNA binding was analysed by electrophoresis on non-denaturing 5% polyacrylamide gels, and the mobilities of protein-complexed and free oligomers detected by autoradiography are indicated.

```
MARRPRHSIY SSDEDDEDIE MCDHDYDGLL PKSGKRHLGK TRWTREEDEK LKKLVEQNGT   60
                                                    R1
DDWKVIANYL PNRTDVQCQH RWQKVLNPEL IKGPWTKEED QRVIELVQKY GPKRWSVIAK  120
                                  R2                          Δ73▼
HLKGRIGKQC RERWHNHLNP EVKKTSWTEE EDRIIYQAHK RLGNRWAEIA KLLPGRTDNA  180
                          R3
  Δ70▼      Δ60▼
IKNHWNSTMR RKVEQEGYLQ EPSKASQTPV ATSFQKNNHL MGPGHASPPS QLSPSGQSSV  240
   Δ50▼                                                      Δ45▼
NSEYPYYHIA EAQNISSHVP YPVALHANIV NVPQPAAAAI QRHYNDEDPE KEKRIKELEL  300

LLMSTENELK GQQALPTQNH TCSYPGWHST SIVDQTRPHG DSAPVSCLGE HHATPSLPAD  360
Δ37▼                    Δ32▼
PGSLPEESAS PARCMIVHQG TILDNVKNLL EFAETLQFID SFLNTSSNHE SSGLDAPTLP  420

STPLIGHKLT PCRDQTVKTQ KENSIFRTPA IKRSILESSP RTPTPFKHAL AAQEIKYGPL  480
  Δ20▼
KMLPQTPSHA VEDLQDVIKQ ESDESGIVAE FQESGPPLLK KIKQEVESPT EKSGNFFCSN  540

HWAENSLSTQ LFSQASPVAD APNILTSSVL MTPVSEDEDN VLKAFTVPKN RPLVGPLQPC  600

SGAWEPASCG KTEDQMTASG PARKYVNAFS ARTLVM                            636
```

Fig. 2. Truncation points of deleted c-*myb* proteins. The amino acid sequence of mouse c-*myb* protein is represented using the single letter code. On this sequence are shown by arrows the truncation points of carboxy-terminally deleted proteins MybΔ20 through MybΔ73; these proteins all initiate with methionine at amino acid position 1. The locations of the three imperfect amino acid repeats (R1, R2 and R3) are identified by underlining.

dimerization domain (the so-called leucine zipper) present in, for example, the *fos* and *jun* oncogene products (Kouzarides and Ziff, 1988; Sassone-Corsi *et al.*, 1988;

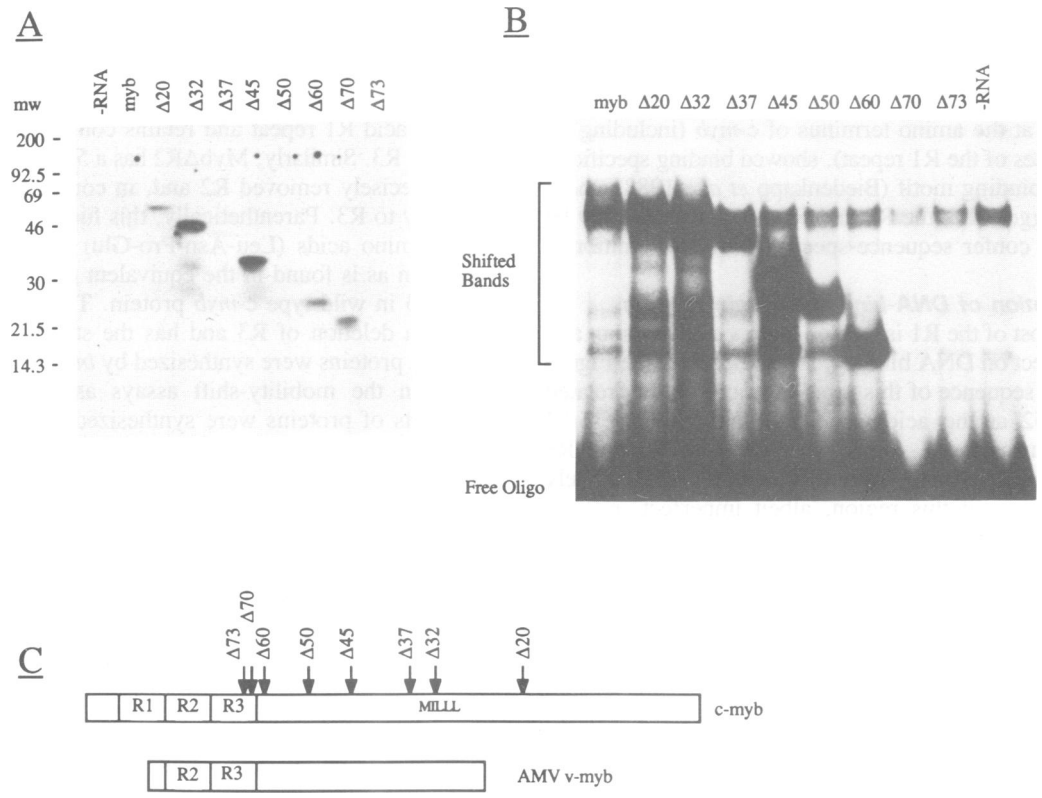

Fig. 3. DNA binding of carboxy-terminally truncated c-*myb* proteins. (**A**) Unlabelled full-length c-*myb* protein and truncated derivatives progressively deleted from the carboxy terminus (MybΔ20 through MybΔ73, see Figure 2) were synthesized by *in vitro* translation and identified by Western blot analysis using the Myb1 antibody. To the left are shown the positions of mol. wt standards run in parallel. (**B**) Reticulocyte lysates containing the proteins characterized above were used in mobility-shift assays with the OHL34/35 DNA oligomer. Protein−DNA complexes were identified by electrophoresis on non-denaturing 5% polyacrylamide gels followed by autoradiography. The track labelled −RNA represents the result obtained with a reticulocyte lysate in which no mRNA was added. (**C**) Locations of the carboxy-terminal truncation points on a diagrammatic representation of the full-length c-*myb* protein. The imperfect amino acid repeats are boxed and the position of the potential leucine zipper (MILLL) is indicated. Below is shown the AMV v-*myb* sequence.

Gentz *et al.*, 1989; Neuberg *et al.*, 1989; Turner and Tjian, 1989).

Truncated proteins were made by *in vitro* translation of synthetic mRNAs transcribed from derivatives of pT7βmyb bearing progressive deletions of 3′ c-*myb* coding sequences. The locations of the protein truncation points, determined by DNA sequencing of the template plasmids, are shown on the amino acid sequence in Figure 2. Recession of amino acids up to and just into the R3 repeat of the reiterated region was achieved in this way, as in mutants pT7βmybΔ70 and pT7βmybΔ73. *In vitro*-translated proteins were tested prior to use in DNA-binding experiments to ensure that products of the appropriate size were made. To this end they were fractionated by SDS−PAGE and detected immunologically using the Myb1 antibody, a polyclonal antibody raised against a bacterially expressed protein comprising approximately the amino-terminal half of mouse c-*myb* (McMahon *et al.*, 1988). It was clear from this anaysis that all the Myb variants were stably produced (Figure 3A) with the apparent exception of that designated MybΔ73, which represents the largest deletion. However, failure to detect this protein probably reflected the absence of epitopes recognized by the antibody, as synthesis of this product was confirmed by labelling proteins with [^{35}S]methionine in parallel *in vitro* translations (data not shown).

The various products were then subjected to the mobility-shift assay, using the same oligomer and conditions as described above. A protein−DNA complex was noted with a number of the truncated products (Figure 3B), and it was apparent that the mobility of these complexes was inversely related to the size of the protein. This supported the conclusion that the experimentally observed mobility shifts resulted from interaction of the oligomer with Myb. Further, it was obvious that removal of carboxy-terminal sequences downstream of amino acid 200, as in MybΔ60, had no discernible effect on DNA binding at the quantitative level (Figure 3B). However, truncation of c-*myb* protein by a further 13 amino acids, as in MybΔ70, resulted in failure to bind DNA in this assay. As MybΔ70 protein lacks just five amino acids of the carboxy terminus of R3 (Figure 2), it appears that violation of this repeat is correlated with the loss of DNA-binding activity. The MybΔ73 protein, which contains a larger deletion of the R3 repeat, also failed to produce a mobility shift of the oligomer (Figure 3B). These results indicate, therefore, that the R3 repeat is a necessary component of a functional DNA-binding domain.

It was also apparent that, apart from MybΔ70 and MybΔ73, the relative amounts of protein−DNA complexes obtained with full-length c-*myb* protein and each of the truncated derivatives approximately reflected their abundance in the reticulocyte lysates as determined by Western blot analysis (Figure 3A and B). Therefore, these data establish that sequences downstream of R3 are not required, either at the qualitative or quantitative level, for specific recognition

of the Myb-binding motif and argue against conserved sequences in this part of the c-*myb* protein (such as the potential leucine zipper) having an essential role in this interaction. Since the v-*myb* protein, which lacks the 71 amino acids at the amino terminus of c-*myb* (including all but 18 residues of the R1 repeat), showed binding specificity to the Myb-binding motif (Biedenkapp *et al.*, 1988), these data also suggest that the R2 and R3 regions alone may be sufficient to confer sequence-specific DNA recognition.

Precise deletion of DNA-binding domain repeats

Although most of the R1 is deleted from v-*myb* without any apparent affect on DNA binding, it is noticeable (see Figure 4A) that the sequence of this repeat is more closely related to R2 (24/52 amino acids identical) than to R3 (16/52 identity). Furthermore, the identity between R2 and R3 (18/52) is less than that between R2 and R1. As it is likely that redundancy of this region, albeit imperfect, reflects reiteration of a functional domain, this raised the possibility that the R1 sequence may be able to substitute for R2 in conferment of DNA-binding specificity. To investigate this possibility, proteins containing two of the three repeats in all possible combinations were generated using the polymerase chain reaction to delete precisely these sequences from pT7βmyb, as illustrated for R2 in Figure 4(B). Thus, the MybΔR1 protein bears a complete deletion of the 52 amino acid R1 repeat and retains complete copies of both R2 and R3. Similarly, MybΔR2 has a 52 amino acid deletion that precisely removed R2 and, in consequence, fused R1 directly to R3. Parenthetically, this fusion placed the same four amino acids (Leu-Asn-Pro-Glu) at the novel R1/R3 junction as is found in the equivalent position between R2 and R3 in wild-type c-*myb* protein. The MybΔR3 protein bears a deletion of R3 and has the structure R1/R2. The variant proteins were synthesized by *in vitro* translation and used in the mobility-shift assays as before: equivalent amounts of proteins were synthesized with each template (data not shown). In the analysis shown (Figure 5), the templates were terminated at a *Sma*I site in the c-*myb* coding sequence, generating proteins that were truncated at amino acid 325 in each case. The aim was to displace the specific protein−DNA complex away from background bands, but equivalent results were obtained with proteins that did not bear such carboxy-terminal deletions (data not shown). We

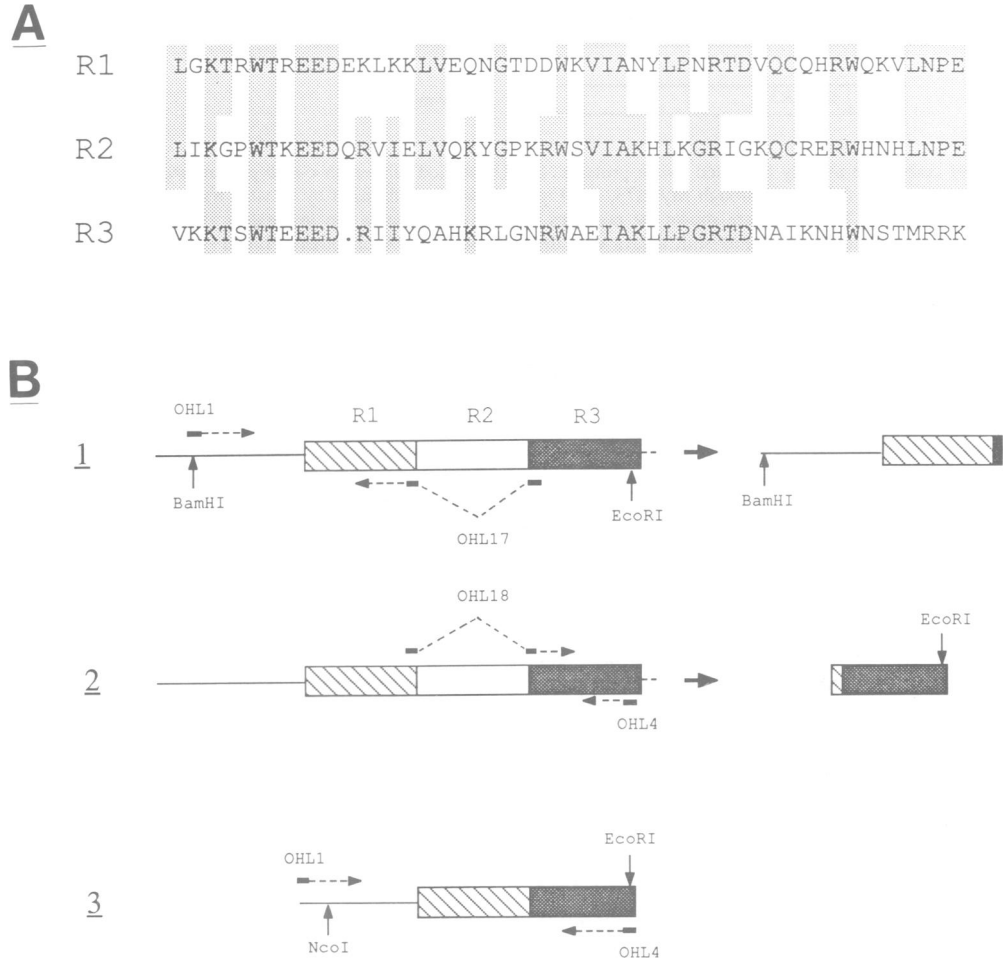

Fig. 4. Deletion of the DNA-binding domain repeats. (**A**) Amino acid sequences of the contiguous imperfect repeats of the mouse c-*myb* DNA-binding domain are represented by the single-letter code. Regions of identity between these sequences are highlighted. (**B**) Strategy for deleting R2 by the polymerase chain reaction is depicted. Reactions 1 and 2 were primed using the single-stranded oligomers indicated (OHL1, OHL4, OHL17 and OHL18). OHL1 and OHL4 contain convenient restriction sites while OHL17 and OHL18 each in effect contain a deletion of nucleotides encoding the R2 sequence. The products of the initial reactions, shown to the right, were mixed and used as a template to generate the ΔR2 product represented in (**3**). The *Nco*I and *Eco*RI restriction sites were utilized to insert the deleted sequence into plasmid pT7βmyb, producing pT7βmybΔR2. Similar strategies were used to delete the R1 and R3 sequences shown in (A) in order to generate pT7βmybΔR1 and pT7βmybΔR3.

found that the DNA-binding activity of MybΔR1 was indistinguishable from that of the wild-type protein (Figure 5), indicating that R1 does not contribute to interactions with the Myb-binding motif. This finding, therefore, supports the supposition that the 18 amino acids of R1 retained by the AMV v-*myb* protein are irrelevant to this interaction. As expected from the previously observed effect of carboxy-terminal truncation of R3 (Figure 3B), the MybΔR3 protein completely lacked specific DNA-binding activity (Figure 5). Moreover, it was found that the MybΔR2 protein also failed to bind the oligomer, despite the sequence similarities between R1 and R2. It appears, then, that R1 is unable to substitute for R2 in generation of a functional DNA-binding domain, suggesting that the roles or specificities of these two repeats are somewhat distinct. These results re-emphasize that a functional DNA-binding domain requires components from both R2 and R3. Sequences outside these repeats have little or no effect on determining this activity, at least as assessed using the assay employed here.

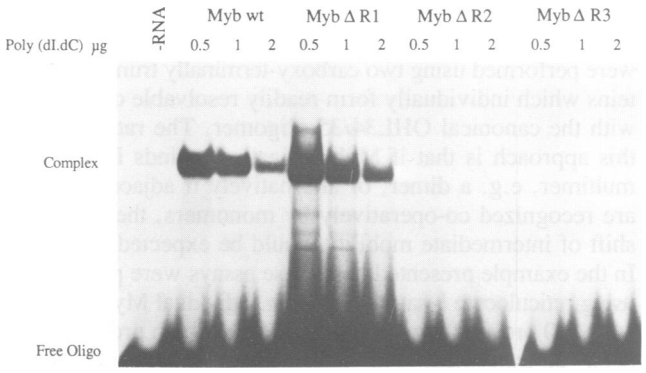

Fig. 5. Mobility-shift analyses with proteins containing precise deletions of DNA-binding domain repeats. Complexes formed between the OHL34/35 DNA oligomer and c-*myb* proteins containing all three repeats (Myb wt) or deletions of R1, R2 or R3 (MybΔR1, MybΔR2 and MybΔR3) were resolved by electrophoresis on non-denaturing 5% polyacrylamide gels and detected by autoradiography. All proteins were truncated at amino acid 325 (see Figure 2) by cutting the mRNA template plasmids at a *Sma*I site. The control track (−RNA) represents the result obtained with a reticulocyte lysate in which no mRNA was added. Binding reactions contained 0.5, 1 or 2 μg of the non-specific competitor poly(dI-dC).

Interaction of c-myb protein with the DNA binding site

It is striking that both the R2 and R3 repeats were required for DNA binding and that the majority of DNA sequences demonstrated to bind Myb (including the OHL34/35 oligomer used in these studies) contain multiple, frequently tandemly arranged Myb-binding motifs (Biedenkapp *et al.*, 1988). One interpretation of this coincidence is that a single protein subunit contacts adjacent Myb-binding motifs through interactions involving both R2 and R3 repeats. If this were

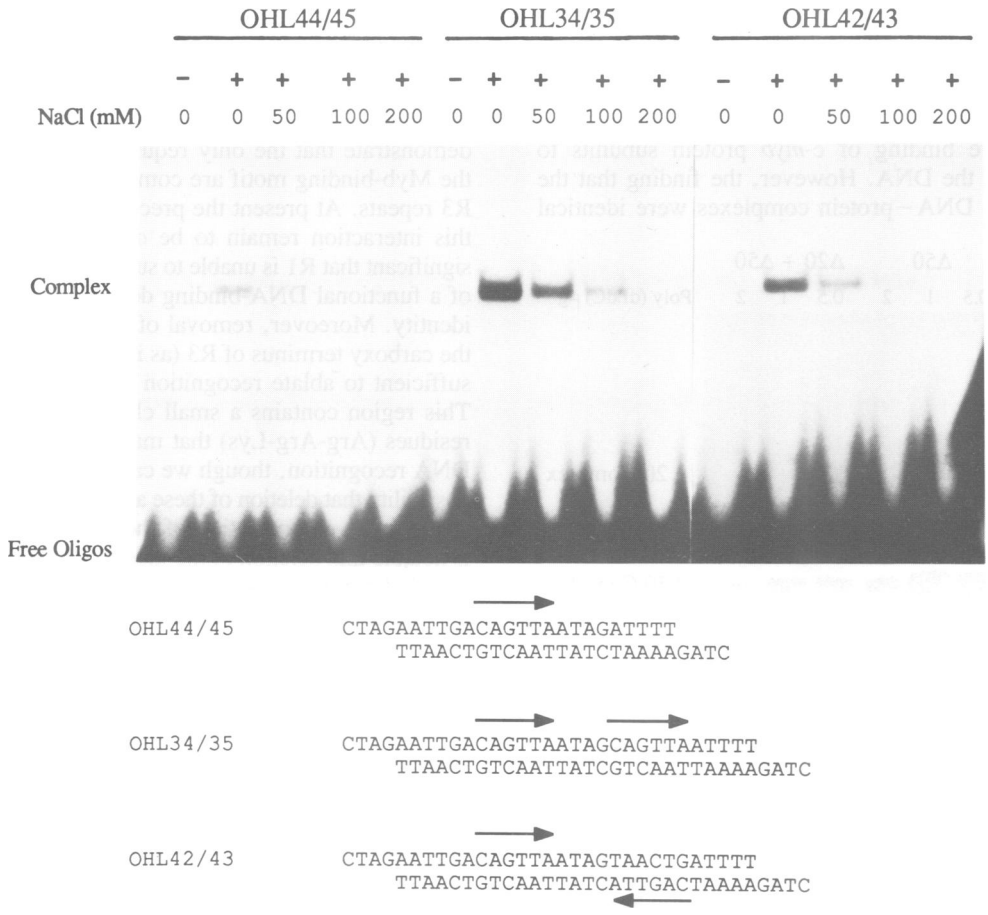

Fig. 6. Mobility-shift analyses with DNA oligomers containing alternative arrangements of the Myb-binding motif. Assays were performed with the indicated ^{32}P-labelled DNA oligomers. Binding reactions included either no lysate (−) or reticulocyte lysates (+) containing the MybΔ45 protein and were performed with 0−200 mM NaCl. Protein−DNA complexes were identified by electrophoresis on non-denaturing 5% polyacrylamide gels followed by autoradiography. The sequences of the oligomers are shown below, on which the relative orientation of the Myb-binding motifs are indicated by arrows.

the case, oligomers containing only one Myb-binding motif or adjacent motifs arranged in opposing orientation, might be expected to bind Myb rather less avidly. To investigate this possibility, further mobility-shift experiments were performed with oligomers that were identical in sequence to OHL34/35 used in previous analyses, other than containing only a single Myb-binding motif or duplicated hexamers arranged in inverse orientation (OHL44/45 and OHL42/43 respectively; Figure 6). All three oligomers were ^{32}P-labelled to the same specific activity so that the amount of complex formation with c-*myb* protein could be directly compared. It was found that with each oligomer a mobility shift was obtained (Figure 6); however, the interactions differed quantitatively. The OHL34/35 oligomer (containing Myb-binding motifs arranged in tandem) gave the greatest amount of complex formation while that with a single motif (OHL44/45) gave comparatively little. The amount of complex obtained with the inverse orientation of motifs (OHL42/43) was intermediate. All interactions were equally sensitive to salt concentration, and the mobility shifts in each instance were qualitatively identical (Figure 6). These results are therefore consistent with a model in which the Myb subunit contacts adjacent Myb-binding motifs on DNA, although it is apparent that some binding can be achieved with a single motif. Perhaps then, specific binding to one Myb-binding motif stabilizes interactions with other non-specific adjacent sequences. The finding that complex formation with inversely oriented motifs was more efficient than with a single motif may simply be explained by the 2-fold higher density of potential binding sites in the latter oligomer.

The data obtained here are also superficially consistent with co-operative binding of c-*myb* protein subunits to adjacent sites on the DNA. However, the finding that the mobilities of the DNA−protein complexes were identical in each case, regardless of whether the target oligomer contained single or duplicate sites, argues against this possibility and suggests that these interactions involve a single Myb subunit. Nevertheless, formally to rule out co-operativity and to provide information on the subunit structure of the bound protein, further mobility-shift assays were performed using two carboxy-terminally truncated proteins which individually form readily resolvable complexes with the canonical OHL34/35 oligomer. The rationale for this approach is that if Myb obligatorily binds DNA as a multimer, e.g. a dimer, or alternatively if adjacent motifs are recognized co-operatively by monomers, then a band-shift of intermediate mobility would be expected to result. In the example presented here these assays were performed using reticulocyte lysates containing individual MybΔ20 and MybΔ50 proteins or with lysates in which both products had been co-translated. No intermediate band-shifts were observed (Figure 7), however, indicating that each oligomer bound only one or other of the truncated proteins. Similar data were obtained using MybΔ20 in conjunction with MybΔ45 and MybΔ45 together with MybΔ50 (data not shown). These data are thus entirely consistent with a model in which a single monomeric c-*myb* protein binds DNA by straddling juxtaposed recognition sites. It is conceivable, but not proven by this analysis, that this recognition involves interaction of R2 and R3 with separate adjacent Myb-binding motifs.

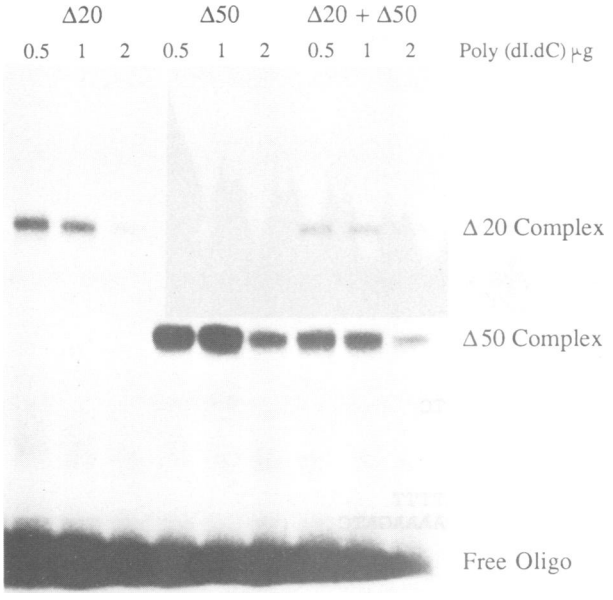

Fig. 7. Mobility-shift analyses with co-translated proteins. Assays were performed using the OHL34/35 DNA oligomer and either MybΔ20 and MybΔ50 proteins translated *in vitro* alone or co-translated together. Binding reactions contained 0.5, 1 or 2 μg of the non-specific competitor poly(dI-dC). Protein−DNA complexes were identified by electrophoresis on non-denaturing 5% polyacrylamide gels and autoradiography.

Discussion

The results presented here, together with previously published findings for v-*myb* (Biedenkapp *et al.*, 1988), demonstrate that the only requirements for recognition of the Myb-binding motif are components of both the R2 and R3 repeats. At present the precise amino acids involved in this interaction remain to be determined. However, it is significant that R1 is unable to substitute for R2 in generation of a functional DNA-binding domain despite sharing 46% identity. Moreover, removal of just five amino acids from the carboxy terminus of R3 (as in MybΔ70, Figure 3B) was sufficient to ablate recognition of the Myb-binding motif. This region contains a small cluster of positively-charged residues (Arg-Arg-Lys) that may be directly implicated in DNA recognition, though we cannot at present exclude the possibility that deletion of these amino acids simply adversely changes the conformation of the DNA-binding domain. It is notable that deletion of the identical five amino acids from v-*myb* did not appear to affect the ability of this protein to interact non-specifically with DNA (Klempnauer and Sippel, 1987), nor indeed did removal of 27 amino-terminal amino acids of R2. It needs to be stressed, however, that non-specific DNA binding may depend primarily on electrostatic interactions. In contrast, the structural features required to recognize the Myb-binding motif probably extend beyond the previously reported confines of the v-*myb* DNA-binding domain. A salient example is provided by the *fos* oncogene; although c-*fos* proteins have been shown to bind non-specifically to DNA−cellulose columns in the absence of associated proteins (Sambucetti and Curran, 1986), specific binding to the AP-1 binding motif is not observed (Gentz *et al.*, 1989). Indeed, such interaction is dependent upon the formation of a DNA-binding domain by dimerization with *jun* oncogene products.

The notion that elements of the Myb repeats contribute to binding specificity is also consistent with the remarkable extent of sequence conservation. Thus, the R2 and R3 repeats of mouse c-*myb* are, respectively, 71 and 84% identical with those of the *Drosophila* homologue (Katzen *et al.*, 1985; Peters *et al.*, 1987). These sequences are conserved to an even greater extent between human c-*myb* and two related human genes, A-*myb* and B-*myb* (Nomura *et al.*, 1988). The extent of this homology strongly suggests that each of these proteins should recognize the same DNA-binding motif. In this respect, it is interesting that expression of all three related human *myb* genes was observed in certain haemopoietic cells (Nomura *et al.*, 1988). It seems likely that the activity of c-*myb* within these cells will be influenced by competition or co-operation between products of this gene family.

The R1 repeat is noticeably less well conserved between species than R2 and R3; the mouse and *Drosophila* R1 repeats are only 48% homologous. None the less, in the context of the entire gene this is a very significant level of conservation as there is little identity outside the reiterated region (Peters *et al.*, 1987). Such a significant level of conservation obviously implies retention of function, but from data presented here and in previous reports (Biedenkapp *et al.*, 1988) it is clear that the R1 repeat is not required for recognition of the Myb-binding motif *in vitro*. That this sequence is subject to extensive truncation in both v-*myb*-containing avian retroviruses (Klempnauer *et al.*, 1983; Nunn *et al.*, 1984) and in most retrovirus-activated c-*myb* genes in murine myeloid leukaemias (Lavu and Reddy, 1986; Shen-Ong *et al.*, 1986) suggests that it is also not an absolute requirement for specific DNA binding *in vivo*. This raises the question of the possible role of R1. On the basis of its homology with R2 and R3, it seems almost inconceivable that it contributes nothing to DNA-binding activity. One possibility is that R1 acts to modify the *in vivo* DNA-binding properties of Myb, e.g. by interacting with additional sequences that lie in close apposition to the Myb-binding motif. This may impose additional constraints on the interaction of c-*myb* protein with DNA and restrict the number of sites that enable binding with high affinity. Truncation of R1-coding sequences in retrovirus-mediated myeloid tumours may remove such constraints. In this regard, it is interesting to note that insertional activation of c-*myb* in avian B cell lymphomas (as opposed to myeloid tumours) appears to be associated with expression of novel proteins in which the amino-terminal truncation is far less extensive (Kanter *et al.*, 1988; Pizer and Humphries, 1989) and in which the entire reiterated DNA-binding region would remain intact. Thus, it is conceivable that these two modes of activation in tumours of distinct haemopoietic lineages reflect requirements for binding to different spectra of target sites in these cells, with specificity perhaps endued by the presence or absence of R1 in the mutated proteins.

The data presented here indicate that c-*myb* protein is able to bind DNA as a monomer and, indeed, the *in vitro*-translated product appears to interact with the target oligomer only in this form (see Figure 7). Other data are also consistent with this conclusion. Thus, we obtained no evidence from immunoprecipitation experiments that stable multimers are formed *in vitro* (unpublished observations). Sequential recession of the carboxy terminus of c-*myb* resulted in an increase in electrophoretic mobility of the resultant protein–DNA complexes in line with the decrease in size of the protein. Moreover, the lack of abrupt and otherwise inexplicable changes in mobility argues against the action of a dimerization domain, at least located in the region of the protein downstream of the reiterated sequence. Of course we cannot exclude the possibility that homo- or hetero-multimeric interactions do occur under other circumstances and that these may contribute to the biochemical activities of Myb *in vivo*. Our conclusion differs from that of Bading (1988) who, on the basis of the electrophoretic mobilities of two complexes formed between purified v-*myb* protein and an 81 bp bacteriophage λ DNA fragment (part of which is identical to the OHL34/35 oligomer used here), proposed that interactions occurred both with protein monomers and dimers. Formation of the monomeric protein–DNA complex was purportedly sensitive to inhibition by poly(dI-dC) at concentrations ~100-fold less than those routinely employed in our studies. However, even in the absence of this non-specific competitor we failed to observe a higher mobility complex with *in vitro*-translated protein that would be consistent with these other findings (data not shown). Although the reasons for this discrepancy remain obscure, it may be significant that the 81 bp λ fragment contains an additional near-consensus Myb-binding motif (CAACTT) that is absent from the oligomer used here. This site might interact separately with Myb, potentially yielding a complex of somewhat different stability and mobility than that formed with the Myb-binding motifs implicated here.

A full appreciation of the relevance of DNA-binding activity on functions determined by either v-*myb* or c-*myb* has yet to emerge. It is of interest, however, that the lesion in a mutant of E26 virus, which is temperature sensitive for transformation of myelomonocytic cells, was localized to the R3 repeat (Frykberg *et al.*, 1988) and resulted in expression of a protein that was temperature sensitive for non-specific DNA binding *in vitro* (Moelling *et al.*, 1985). Further evidence for the requirement of DNA binding in transformation by v-*myb* has been obtained from an analysis of virus mutants bearing deletions within this gene (Ibanez *et al.*, 1988): while DNA-binding activity appears to be necessary for transformation it does not seem to be sufficient, and it must be presumed that other functions expressed by this protein are also vital. One such additional function might be the ability to transactivate transcription from certain promoters, e.g. promoters under the influence of the SV40 enhancer (Nishina *et al.*, 1989). It may be relevant that this enhancer contains a single copy of a consensus Myb-binding motif (CAGTTA) located between the GTII and GTI motifs (Zenke *et al.*, 1986) and further that transactivation was observed only when the enhancer sequence was present in multiple copies. Based on our observation that oligomers with tandem hexamer motifs were more efficient in forming a complex with Myb than those containing a single motif (Figure 6), one might speculate that the R2 and R3 repeats act in this instance by binding to separate sites in adjacent SV40 enhancers, perhaps bringing together two sequences that may otherwise lie some distance apart. This would also have the effect of looping out the DNA between the Myb-binding motifs. Either or both events may be significant in transcriptional activation. We intend to study the effects on complex formation of increasing the distance between adjacent Myb-binding motifs to determine whether the findings are consistent with our current conception of this Myb–DNA interaction.

Materials and methods

In vitro transcription

To generate c-*myb* transcripts *in vitro*, a 3.05 kb mouse c-*myb* cDNA fragment delimited by *Nco*I and *Eco*RI sites (Watson *et al.*, 1987) was first inserted into the vector pT7βΔ6Sal (Norman *et al.*, 1988) using the equivalent restriction sites. This placed the c-*myb* initiation codon (which lies within the 5′-most *Nco*I recognition site) and the remainder of the coding sequence immediately downstream of a β-globin 5′ untranslated leader, which in turn is located adjacent to a T7 RNA polymerase promoter (derived from the pGEM-2 vector; Promega Biotec). Transcription of a defined fragment containing the entire c-*myb* coding sequence was achieved by cutting the resultant pT7βmyb plasmid at a unique *Bgl*II site located within the 3′ untranslated cDNA region. This template (2 μg) was incubated at 37°C for 1 h in a total volume of 100 μl containing 40 mM Tris−HCl, pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 0.5 mM each rNTP, 100 U RNasin (Promega Biotec) and 50 U T7 RNA polymerase (Promega Biotec). The RNA product (the typical yield was 10 μg) was phenol/chloroform extracted three times and recovered by ethanol precipitation.

Construction of deletion mutants

Carboxy-terminal coding sequences were deleted by a combination of exonuclease III and mung bean nuclease digestion using a kit and conditions supplied by Stratagene. Following deletion, fragments were recloned into a derivative of pT7βmyb just upstream from a universal translation terminator (CTAGTTAACTAG). The locations of end-points were determined by DNA sequencing appropriate regions of the double-stranded plasmids using the protocol supplied by Stratagene.

Specific deletion of individual DNA-binding domain repeats was obtained using the polymerase chain reaction as outlined in Figure 4(B). Typically, two reactions were primed initially using pairs of complementary oligomers that each spanned 20 nt both sides of the DNA sequence to be deleted, together with an additional oligomer that encompassed a suitable restriction site on either side of the reiterated region (i.e. *Nco*I and *Eco*RI or *Pst*I). Reactions were done under mineral oil in a total volume of 100 μl containing 50 mM KCl, 10 mM Tris−HCl, pH 8.3, 2.5 mM $MgCl_2$, 0.02% gelatin, 0.2 mM of each dNTP, 0.05 nmol of each primer, 0.5 μg template DNA and 2.5 U Taq DNA polymerase (New England Biolabs). Tubes were placed in a programmable heating block and subjected to 20 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 3 min and elongation at 70°C for 4 min. The final elongation step was extended to 10 min. The products of these reactions were purified by PAGE and eluted as described in Maxam and Gilbert (1980). A second reaction was then carried out using a mixture of the two primary products (~10 ng) and the two restriction site-encompassing oligomers used in the first reactions. After phenol/chloroform extraction, products were digested with restriction enzymes that cut at suitable sites either side of the deletion: the sites used were not necessarily those contained by the flanking oligomers, but were chosen on the basis of the minimum size DNA fragment that could be conveniently recloned. After purification by PAGE, the isolated fragments were cloned into pEMBL vectors (Dente *et al.*, 1983) so that single-stranded DNA could be obtained by bacteriophage f1 infection of transformants. Inserts were completely sequenced by the dideoxy chain termination method using a kit and conditions supplied by United States Biochemicals. It was generally found that approximately half the inserts contained polymerase-generated mutations (e.g. transitions or single base deletions) in addition to the desired deletion. Appropriate unmutated fragments were inserted into pT7βmyb.

In vitro translation

Approximately 0.5 μg of the synthetic RNAs were used to programme *in vitro* translation in a volume of 50 μl in rabbit reticulocyte lysates (obtained from Amersham International plc or Promega Biotec) using conditions specified by the suppliers (Jackson and Hunt, 1983). The ability of each batch of synthetic RNA to direct synthesis of appropriately sized proteins was first tested in this system by [^{35}S]methionine labelling and SDS−PAGE analysis. For use in mobility-shift assays, proteins were not radioactively labelled and were used directly without further dilution or purification. Generally, samples of unlabelled proteins were quantitated by Western blot analysis with the Myb1 polyclonal rabbit antibody using conditions described previously (McMahon *et al.*, 1988).

Mobility-shift assays

Formation of protein −DNA complexes was achieved by incubating 5 μl of reticulocyte lysate with ~250 pg of each double-stranded oligomer (labelled to a specific activity of 4×10^7 c.p.m./μg using [^{32}P]dCTP and the Klenow enzyme) in a final volume of 20 μl. Where protein binding to different oligomers was examined, the specific activities were adjusted after labelling and equivalent molar amounts and radioactivity were used in each assay. Standard binding buffer contained 20 mM HEPES, pH 7.9, 20% glycerol, 5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT and 2 μg of the non-specific inhibitor poly(dI-dC), and the reactions were carried out at 30°C for 30 min. Occasionally these conditions were altered by varying the amount of poly(dI-dC) or by the addition of NaCl at concentrations up to 200 mM as indicated in the figure legends. Following the binding reaction, the entire samples were loaded onto 5% polyacrylamide gels (the acrylamide to bis-acrylamide ratio was 40:1) cast in 0.5 × TBE (1 × TBE is 50 mM Tris, 50 mM boric acid, 1 mM EDTA) and were resolved by electrophoresis at 10 V/cm using the same buffer. Gels were fixed using 10% acetic acid, dried and subjected to autoradiography.

Acknowledgements

We thank Richard Treisman for provision of pT7βΔ6Sal and Gordon Peters and Gerard Evan for critical appraisal of the manuscript.

References

Bading,H. (1988) *Nucleic Acids Res.*, **16**, 5241−5248.
Bading,H., Hansen,J. and Moelling,K. (1987) *Oncogene*, **1**, 395−401.
Bender,T.P. and Kuehl,W.M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3204−3208.
Biedenkapp,H., Borgmeyer,U., Sippel,A. and Klempnauer,K.-H. (1988) *Nature*, **335**, 835−837.
Bishop,J.M. and Varmus,H. (1985) In Weiss,R., Teich,N., Varmus,H. and Coffin,J. (eds), *The Molecular Biology of Tumor Viruses, Part IV: RNA Tumor Viruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 351−356.
Boyle,W.J., Lampert,M.A., Lipsick,J.S. and Baluda,M.A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4265−4269.
Boyle,W.J., Lipsick,J.S. and Baluda,M.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 4685−4689.
Dente,L., Sollazzo,M., Baldari,C., Cesareni,G. and Cortese,R. (1985) In Glover,D.M. (ed.), *DNA Cloning*. IRL Press, Oxford, Vol. 1, p. 101−107.
Frykberg,L., Metz,T., Brady,G., Introna,M., Beug,H., Vennstrom,B. and Graf,T. (1988) *Oncogene Res.*, **3**, 313−322.
Gentz,R., Rauscher,F.J., Abate,C. and Curran,T. (1989) *Science*, **243**, 1695−1699.
Gonda,T.J., Gough,N.M., Dunn,A.R. and de Blaquiere,J. (1985) *EMBO J.*, **4**, 2003−2008.
Ibanez,C.E., Garcia,A., Stober-Grasser,U. and Lipsick,J.S. (1988) *J. Virol.*, **62**, 4398−4402.
Jackson,R.J. and Hunt,R.T. (1983) *Methods Enzymol.*, **96**, 50−74.
Kanter,M.R., Smith,R.E. and Hayward,W.S. (1988) *J. Virol.*, **62**, 1423−1432.
Katzen,A.L., Kornberg,T.B. and Bishop,J.M. (1985) *Cell*, **41**, 449−456.
Klempnauer,K.-H. and Sippel,A.E. (1986) *Mol. Cell. Biol.*, **6**, 62−69.
Klempnauer,K.-H. and Sippel,A.E. (1987) *EMBO J.*, **6**, 2719−2725.
Klempnauer,K.-H., Ramsay,G., Bishop,J.M., Moscovici,G.M., Moscovici,C., McGrath,J.P. and Levinson,A.D. (1983) *Cell*, **33**, 345−355.
Klempnauer,K.-H., Symonds,G., Evan,G.I. and Bishop,J.M. (1984) *Cell*, **37**, 537−547.
Kouzarides,T. and Ziff,E. (1988) *Nature*, **336**, 646−651.
Lavu,S. and Reddy,E.P. (1986) *Nucleic Acids Res.*, **14**, 5309−5320.
Majello,B., Kenyon,L.C. and Dalla-Favera,R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9636−9640.
Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 43−62.
McMahon,J., Howe,K.M. and Watson,R.J. (1988) *Oncogene*, **3**, 717−720.
Moelling,K., Pfaff,E., Beug,H., Beimling,P., Schaller,H.E. and Graf,T. (1985) *Cell*, **40**, 983−990.
Mushinski,J.F., Potter,M., Bauer,S.R. and Reddy,E.P. (1983) *Science*, **220**, 795−798.
Neuberg,M., Schuermann,M., Hunter,J.B. and Muller,R. (1989) *Nature*, **338**, 589−590.
Nishina,Y., Nakagoshi,H., Imamoto,F., Gonda,T. and Ishii,S. (1989) *Nucleic Acids Res.*, **17**, 107−116.
Norman,C., Runswick,M., Pollock,R. and Treisman,R. (1988) *Cell*, **55**, 989−1003.
Nomura,N., Takahashi,M., Matsui,M., Ishii,S., Date,T., Sasamoto,S. and Ishizaki,R. (1988) *Nucleic Acids Res.*, **16**, 11075−11089.

Nunn,M., Weiher,M., Bullock,P. and Duesberg,P. (1984) *Virology*, **139**, 330–339.
Paz-Ares,J., Ghosal,D., Wienand,U., Peterson,P.A. and Saedler,H. (1987) *EMBO J.*, **6**, 3553–3558.
Peters,C.W.B., Sippel.A.E., Vingron,M. and Klempnauer,K.-H. (1987) *EMBO J.*, **6**, 3085–3090.
Pizer,E. and Humphries,E.H. (1989) *J. Virol.*, **63**, 1630–1640.
Rosson,D. and Reddy,E.P. (1986) *Nature*, **319**, 604–606.
Sambucetti,L.C. and Curran,T. (1986) *Science*, **234**, 1417–1419.
Sassone-Corsi,P., Ransone,L.J., Lamph,W.W. and Verma,I.M. (1988) *Nature*, **336**, 692–695.
Shen-Ong,G.L.C., Morse,H.C., Potter,M. and Mushinski,J.F. (1986) *Mol. Cell. Biol.*, **6**, 380–392.
Slamon,D.J., Boone,T.C., Murdock,D.C., Keith,D.E., Press,M.F., Larson,R.A. and Souza,L.M. (1986) *Science*, **233**, 347–351.
Turner,R. and Tjian,R. (1989) *Nature*, **243**, 1689–1694.
Watson,R.J., Dyson,P.J. and McMahon,J. (1987) *EMBO J.*, **6**, 1643–1651.
Weinstein,Y., Ihle,J.N., Lavu,S. and Reddy,E.P. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5010–5014.
Zenke,M., Grundstrom,T., Matthes,H., Wintzerith,M., Schatz,C., Wildeman,A. and Chambon,P. (1986) *EMBO J.*, **5**, 387–397.

Received on September 6, 1989; revised on November 2, 1989